Vibrational Spectra of Phosphate Ions in Aqueous Solution Probed by First-Principles Molecular Dynamics

VandeVondele, Joost ; Troester, Philipp ; Tavan, Paul ; Mathias, Gerald

Abstract: We have carried out “first-principles” Born-Oppenheimer molecular dynamics (BOMD) simulations of the phosphate ions H2PO4- and HPO42- in liquid water and have calculated their IR spectra by Fourier transform techniques from the trajectories. IR bands were assigned by a so-called “generalized normal coordinate analysis”. The effects of including Hartree-Fock (HF) exchange into the density functional theory (DFT) computation of forces were studied by comparing results obtained with the well-known BP, BLYP, and B3LYP functionals. The neglect of dispersion in the functionals was empirically corrected. The inclusion of HF exchange turned out to yield dramatically improved and, thus, quite accurate descriptions of the IR spectra observed for H2PO4- and HPO42- in aqueous solution. An analysis of earlier computational results (Klahn, M. et al. J. Phys. Chem. A 2004, 108, 6186-6194) on these vibrational spectra, which had been obtained in a hybrid setting combining a BP description of the respective phosphate with a simple molecular mechanics (MM) model of its aqueous environment, revealed three different sources of error, (i) the BP force field of the phosphates is much too soft and would have required a substantial scaling of frequencies, (ii) the oversimplified water force field entailed incorrect solvation structures and, thus, qualitatively wrong patterns of solvatochromic band shifts, and (iii) quantitative frequency computations additionally required the inclusion of HF exchange. Thus, the results of the B3LYP BOMD simulations do not only characterize physical properties like the IR spectra or the solvation structures of the phosphate systems but also provide clues for the future design of simplified but nevertheless reasonably accurate DFT/MM methods applicable to phosphates.

DOI: https://doi.org/10.1021/jp211783z

Vibrational Spectra of Phosphate Ions in Aqueous Solution Probed by First Principles Molecular Dynamics

Joost VandeVondele,[†,¶] Philipp Tröster,[‡] Paul Tavan,[‡] and Gerald Mathias[*,‡]

Institut für Physikalische Chemie, Universität Zürich, and

Lehrstuhl für Biomolekulare Optik, Ludwig-Maximilians-Universität München

E-mail: gerald.mathias@physik.uni-muenchen.de

[*]To whom correspondence should be addressed

[†]Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

[‡]Ludwig-Maximilians-Universität, Oettingenstr. 67, 80538 München, Germany

[¶]current address: ETH Zürich, Wolfgang-Pauli-Strasse 10, 8093 Zurich, Switzerland

KEYWORDS: solvatochromy, hybrid density functional, time correlation, photo-electron spectroscopy, charge transfer

Abstract

We have carried out "first principles" Born-Oppenheimer molecular dynamics (BOMD) simulations of the phosphate ions $H_2PO_4^-$ and HPO_4^{2-} in liquid water and have calculated their IR spectra by Fourier transform techniques from the trajectories. IR bands were assigned by a so-called "generalized normal coordinate analysis". The effects of including Hartree-Fock (HF) exchange into the density functional theory (DFT) computation of forces were studied by comparing results obtained with the well-known BP, BLYP, and B3LYP functionals, respectively. The neglect of dispersion in the functionals was empirically corrected. The inclusion of HF exchange turned out to yield dramatically improved and, thus, quite accurate descriptions of the IR spectra observed for $H_2PO_4^-$ and HPO_4^{2-} in aqueous solution. An analysis of earlier computational results (Klähn *et al. J. Phys. Chem. A* **2004**, *108*, 6186-6194) on these vibrational spectra, which had been obtained in a hybrid setting combining a BP description of the respective phosphate with a simple molecular mechanics model (MM) of its aqueous environment, revealed three different sources of error: (i) the BP force field of the phosphates is much too soft and would have required a substantial scaling of frequencies, (ii) the oversimplified water force field entailed incorrect solvation structures and, thus, qualitatively wrong patterns of solvatochromic band shifts, and (iii) quantitative frequency computations additionally require the inclusion of HF exchange. Thus, the results of the B3LYP BOMD simulations do not only characterize physical properties like the IR spectra or the solvation structures of the phosphate systems, but also provide clues for the future design of simplified but nevertheless reasonably accurate DFT/MM methods applicable to phosphates.

Introduction

Phosphates are key building blocks of phospholipids, DNA, RNA, and of the various nucleotides involved in bioenergetics (e.g. adenosine triphosphate, ATP) and cellular signaling (e.g. guanosine

triphosphate, GTP). Particularly the enzymatically catalyzed hydrolysis of ATP drives a plethora of energy consuming biological processes, whereas the hydrolysis of GTP usually triggers conformational changes of the catalyzing enzymes, the so-called G-proteins, whose modified conformations then represent specific biochemical signals.[1]

Vibrational spectroscopy is a versatile technique to monitor such reactions.[2–8] If one could decode such spectra in terms of the underlying molecular structures and interactions, one additionally might be able to pin down the detailed mechanisms of these reactions. Hybrid methods, which combine a density functional theory (DFT) treatment of a molecule with a simplified molecular mechanics (MM) description of its condensed phase environment, should be well-suited for such a theoretical analysis.[9] Therefore, DFT/MM techniques have been applied to compute the condensed phase vibrational spectra of phosphates.[10–12] In these calculations the IR spectra were derived through the specific protocols of "instantaneous normal mode analysis" (INMA), which have been originally suggested by Nonella *et al.*[13] and further refined by Schmitz and Tavan.[14,15]

However, the infrared (IR) spectra computed for the phosphates[10–12] did not quite reach the high quality found in many other DFT/MM applications.[9,13,16–20] Particularly the solvatochromic shifts calculated by DFT/MM INMA for the vibrational bands of the phosphate ions HPO_4^{2-} and $H_2PO_4^-$ seemed to be too small.[10] These apparent underestimates were attributed to shortcomings of the MM force field used for the aqueous solvent.[9,10] Here, the water molecules had been described by the simple "three point transferable potential"[21] (TIP3P), which lacks polarizability and, due to its planar distribution of atomic partial charges, features an erroneous quadrupole moment. As a result, one may expect that this MM water model does not correctly sample the the first solvation shell of the phosphate ions, which could explain the underestimates of the solvatochromic shifts.

As of today, the actual reasons for the unsatisfactory performance of DFT/MM INMA on the condensed phase vibrational spectra of phosphates still remain elusive. Beside the mentioned shortcomings of the TIP3P water model alternative explanations are conceivable. For instance, using the generalized gradient approximation (GGA) in the form of the gradient-corrected exchange func-

tional of Becke,[22] which had been combined with the correlation functional of Perdew[23] (BP) for the DFT description of the phosphate anions embedded in liquid MM water,[10] could have been a sub-optimal choice. Here, the inclusion of Hartree-Fock (HF) exchange[24] into the popular Becke-Lee-Yang-Parr (BLYP) approach,[22,25] which then is commonly denoted as B3LYP, could lead to better results.[26] Thus the specific technical question here is whether exchanging the above BP/MM approach to the condensed phase IR spectra of phosphates by B3LYP/MM can yield substantially improved descriptions.

Because phosphates play a prominent role in biochemistry and because their reactions can be monitored by vibrational spectroscopy, an understanding of the shortcomings indicated above and the subsequent development of quantitatively more reliable DFT/MM descriptions remain important goals. Such methods should then represent suitable means to compute the finger prints of enzymatic reactions in vibrational spectra, which monitor the hydrolysis of nucleotides in G-proteins and in bioenergetics.

For an improved understanding of the phosphate vibrational spectra and the associated solvatochromic shifts here we compute the IR spectra of the phosphate ions $H_2PO_4^-$ and HPO_4^{2-} in aqueous solution from quite extended "first principles" Born-Oppenheimer molecular dynamics (BOMD) simulations. To check the influence of HF exchange, we go beyond the GGA by employing the B3LYP functional in addition to BLYP despite the strongly enhanced computational effort associated with this extension. According to the analysis of Neugebauer and Hess[26] the inclusion of HF exchange should yield strongly improved descriptions of vibrational spectra, particularly if anharmonicities are properly considered. Note in this context that the use of a dynamics based approach toward the computation of the IR spectra[13,15,27–39] can, to some extent, avoid the pitfalls of the harmonic approximation, which was applied in the BP/MM INMA computations quoted above. The results of our BOMD simulations are expected to provide important insights and reference data that are needed for the construction of improved DFT/MM methods which are capable of accurately describing the IR spectra of phosphates in condensed phase.

By comparing the results of the BLYP BOMD simulations, which apply the GGA, with the

predictions of the B3LYP BOMD simulations and with experimental data[10] on the IR spectra of the phosphate ions $H_2PO_4^-$ and HPO_4^{2-} in solution we want to find out, whether the inclusion of HF exchange actually leads to substantially improved descriptions. Such an analysis can show whether the unsatisfactory performance of the previous BP/MM INMA calculations[10] is mainly caused by the GGA or by the simple TIP3P model of the liquid water environment. Beyond an improved understanding of the phosphate vibrational spectra, analyzing the virtual realities generated by the BLYP- and B3LYP BOMD simulations can also yield insights into the solvation of the phosphate anions by water and into other properties of these solute-solvent systems.

Methods

For the first principles simulations of the phosphate ions in aqueous solution we set up two small periodic simulation systems each containing 64 water molecules and one HPO_4^{2-} or $H_2PO_4^-$ anion, respectively. They were constructed in an MM setting by the following procedure.

Construction of simulation systems. To determine proper volumes for the BOMD simulation systems, first two cubic boxes of 4 nm side lengths were filled with a total of 2190 TIP3P water molecules.[21] Each box additionally contained a single phosphate ion modeled by the respective MM force field given in Ref. 10. Using toroidal boundary conditions as implemented in the MD simulation program EGO for the long-range electrostatics[40] and Berendsen thermo- and barostats[41] the systems were equilibrated by MM-MD for 10 ns at the temperature $T = 295\,\mathrm{K}$ and the pressure $p = 1\,\mathrm{bar}$. 2000 snapshots were drawn from the last 2 ns and the average numbers of water oxygens $n(L)$, which were enclosed by cubic boxes of varying box lengths L centered around the phosphate atoms, were determined. The desired system size $n = 64$ was reached at $L = 1.24\,\mathrm{nm}$ for $H_2PO_4^-$ and at $L = 1.26\,\mathrm{nm}$ for HPO_4^{2-}. Two cubic periodic simulation systems of the thus obtained small sizes were cut out of the two large MM-MD simulation systems, respectively. After energy minimization these periodic phosphate/water configurations provided the initial conditions for our subsequent Born-Oppenheimer MD simulations (BOMD) within the framework of Kohn-Sham

DFT.

DFT and BOMD settings. The BOMD simulations were performed with the CP2K program package.[42] The core charges and electrons were treated by the pseudopotentials proposed by Goedecker, Teter, and Hutter.[43,44] The Kohn-Sham orbitals were expanded in a triple zeta valence basis set including double polarization functions (TZV2P) for all atoms.[42] In order to efficiently compute the electrostatic interactions, the Gaussian and plane waves (GPW) scheme[45] as implemented in the QUICKSTEP module[42] of CP2K was employed. The electron density was expanded in a plane wave basis using a 400 Ry cutoff.

The GPW scheme allows for a linear scaling calculation of the Kohn-Sham matrix, if functionals applying the GGA are employed. Calculations with hybrid functionals are computationally much more demanding, especially for condensed phase systems and in combination with *ab-initio* MD. However, using a combination of efficient screening, of in-core compression, and of a truncated Coulomb operator, the feasibility of performing robust and efficient *ab-initio* MD based on hybrid functionals has recently been demonstrated.[46,47]

For the exchange we employed a screening threshold of 10^{-7} Hartree and a Coulomb operator truncation radius of 6 Å. Furthermore, we applied the auxiliary density matrix method, which further reduces the cost of the HF exchange computation.[48] The auxiliary density matrix was obtained from fitting the wave function to a so-called pFIT3 basis.[48] Subsequently, the resulting density matrix was purified.[48]

As representatives for the classes of GGA and hybrid functionals we chose the BLYP[22,49] and B3LYP[50,51] functionals, respectively. Both functionals were corrected for the missing dispersion interactions using the empirical scheme suggested by Grimme[52] resulting in descriptions called BLYP-D and B3LYP-D. Note here that BLYP-D model with the TZV2P basis set provides a surprisingly[53] reasonable structure and the proper density for liquid water.[54]

For the BOMD simulations we chose a timestep of 0.5 fs. The self-consistent field (SCF) procedure applied the orbital transformation method described in Ref. 55. The SCF convergence

criterion was set to 10^{-7}.

BOMD simulations. The thermal motion in each of the two periodic water/phosphate boxes described further above was simulated by BOMD using the BLYP-D and the B3LYP-D functional. Here, the four simulation systems were first equilibrated in the canonical ensemble for 4 ps at $T =$ 300 K by employing massive Nosé–Hoover chain thermostats.[56] Subsequently, MD trajectories were computed in the micro-canonical (NVE) ensemble, covering 54 ps and 56 ps for $HPO_4^{2-}$and $H_2PO_4^-$, respectively, while using the the BLYP functional. For the B3LYP-D simulation systems we computed two trajectories for each solvated species spanning 33 ps and 22 ps for $H_2PO_4^-$ as well as 34 ps and 28 ps for HPO_4^{2-}. Here, the initial conditions for the second set of trajectories were generated by short (5 ps) intermittent equilibrations in the canonical NVT ensemble, which served to ensure mutual statistical independence.

Coordinates, Mulliken charges, and Wannier centers[57] were extracted every 2 fs from the various trajectories for subsequent statistical analyses. From these data we evaluated at all time points t and for all molecules in the systems molecular charges $Q_i(t)$ and dipole moments $\boldsymbol{\mu}_i(t)$. The $Q_i(t)$ were obtained from the Mulliken charges. The $\boldsymbol{\mu}_i(t)$ were calculated from the locations and charges of the Wannier centers and ion cores using the respective centers of mass as local reference points. The total dipole moments $\boldsymbol{\mu}(t)$ of the simulation systems were evaluated by summing over the molecular dipoles $\boldsymbol{\mu}_i(t)$.

Computation and analysis of IR spectra. Approximating the trajectories of the dipole operator by those of the dipole moment $\boldsymbol{\mu}(t)$ the IR absorption cross sections

$$\alpha(\omega) = \frac{2\pi\beta\omega^2}{3Vc_0 n(\omega)} \int_{-\infty}^{+\infty} M(t)\, e^{-i\omega t} dt. \tag{1}$$

of the simulation systems are the Fourier transforms (FTs) of the ensemble averaged autocorrelation functions (ACFs)

$$M(t) = \langle \boldsymbol{\mu}(0) \cdot \boldsymbol{\mu}(t) \rangle \tag{2}$$

of the dipole moments $\boldsymbol{\mu}(t)$.[14,36,58,59] In Eq. (1), V is the sample volume, $\beta = 1/k_B T$ the inverse temperature, $n(\omega)$ the refractive index, c_0 the speed of light, and the so-called harmonic quantum correction factor[14,59–61]

$$\mathrm{QC}(\omega) = \frac{\omega}{1 - e^{-\beta\hbar\omega}} \tag{3}$$

is included.

Because simulations are restricted to trajectories of finite duration, the ensemble average $M(t)$ has to be replaced by a statistical estimate $M_\tau(t)$ referring to a finite time scale τ. Following the suggestion in Ref. 36 a suitable estimate is given by

$$M_\tau(t) = \int_{-\infty}^{\infty} w_\tau(t'|t)\boldsymbol{\mu}(t') \cdot \boldsymbol{\mu}(t'+t)dt'. \tag{4}$$

Here, the normalized weighting function

$$w_\tau(t'|t) = \frac{g_\tau(t')g_\tau(t'+t)}{g_{\sqrt{2}\tau}(t)} \tag{5}$$

is defined in terms of the normalized Gaussian

$$g_\tau(t) = \frac{1}{\sqrt{2\pi}\tau} \exp\left(\frac{-t^2}{2\tau^2}\right) \tag{6}$$

of width τ.

$M_\tau(t)$ is used to estimate the IR cross-section $\alpha(\omega)$ of a simulation system as given by Eq. (1) through FT as follows.[36] First the dipole moments $\boldsymbol{\mu}(t)$ collected from a given trajectory of duration d and centered around $t = 0$ are multiplied with the Gaussian window function Eq. (6) yielding the modified trajectory

$$\boldsymbol{\mu}_g(t) = g_\tau(t)\boldsymbol{\mu}(t) \tag{7}$$

of weighted dipole moments. Combining Eqs. (4), (5), and (7) yields

$$g_{\sqrt{2}\tau}(t)M_\tau(t) = \int_{-\infty}^{\infty} \boldsymbol{\mu}_g(t') \cdot \boldsymbol{\mu}_g(t'+t)dt'. \tag{8}$$

Thus, the estimated dipole ACF $M_\tau(t)$ weighted by a somewhat wider Gaussian window $g_{\sqrt{2}\tau}(t)$ is the ACF of the weighted dipole trajectory $\boldsymbol{\mu}_g(t)$.

The FT of Eq. (8) follows from the convolution theorem. Because the l.h.s. is the product $g_{\sqrt{2}\tau}(t)M_\tau(t)$ of two functions, their FT is a convolution $\hat{g}_{\sqrt{2}\tau}(\omega) * \hat{M}_\tau(\omega)$ of their respective FTs (denoted by hats). Similarly, because the r.h.s. is, up to sign, a convolution of two functions, their FT is a product. Using the relation $\hat{g}_{\sqrt{2}\tau}(\omega) = \omega_s g_{\omega_s}(\omega)$, according to which the FT of a normalized Gaussian of width $\sqrt{2}\tau$ is essentially again a normalized Gaussian of width $\omega_s = 1/\sqrt{2}\tau$, one gets

$$g_{\omega_s}(\omega) * \hat{M}_\tau(\omega) = \frac{2\pi}{\omega_s} \hat{\boldsymbol{\mu}}_g^*(\omega) \cdot \hat{\boldsymbol{\mu}}_g(\omega). \tag{9}$$

Note that the l.h.s. of Eq. (9) is a smoothed version of the Fourier transformed estimate $M_\tau(t)$ for the dipole ACF $M(t)$. To approximate the IR absorption cross-section $\alpha(\omega)$ we replace the FT $\hat{M}(\omega)$ of $M(t)$ occurring in Eq. (1) by this expression, i.e. by $\hat{M}(\omega) \approx g_{\omega_s}(\omega) * \hat{M}_\tau(\omega)$. With Eq. (9) we finally get for the IR cross-section

$$\alpha(\omega) \approx \frac{4\pi^2\beta\omega^2}{3Vc_0\omega_s n(\omega)} \hat{\boldsymbol{\mu}}_g^*(\omega) \cdot \hat{\boldsymbol{\mu}}_g(\omega). \tag{10}$$

For the computation of IR spectra one thus must solely evaluate the scalar product $\hat{\boldsymbol{\mu}}_g^*(\omega) \cdot \hat{\boldsymbol{\mu}}_g(\omega)$ involving the FT $\hat{\boldsymbol{\mu}}_g(\omega)$ of the weighted dipole moments $\boldsymbol{\mu}_g(t)$ and its complex conjugate $\hat{\boldsymbol{\mu}}_g^*(\omega)$. $\hat{\boldsymbol{\mu}}_g(\omega)$ is efficiently computed using fast FT methods.

Interestingly, the spectral resolution of $\alpha(\omega)$ as approximated by Eq. (10) is hardly diminished as long as wide Gaussian windows $g_\tau(t)$ are employed, since then the kernel $g_{\omega_s}(\omega)$ smoothing $\hat{M}_\tau(\omega)$ is narrow [cf. Eq. (9)]. On the other hand, reducing the width τ of the time windows leads to somewhat wider Gaussian convolution kernels $g_{\omega_s}(\omega)$ and, therefore, suppresses the spectral noise

which inevitably results from limited sampling. In this study we chose τ such that the spectra were smoothed by a Gaussian of width $\omega_s = 10\,\mathrm{cm}^{-1}$.

Extraction of the phosphate contribution to the IR spectra. Because the IR spectra of the solvated phosphate species are of particular interest, the total cross section of the sample was separated into contributions of the phosphates and the water by splitting the total dipole moment

$$\boldsymbol{\mu} = \boldsymbol{\mu}^{\mathrm{p}} + \boldsymbol{\mu}^{\mathrm{w}} \tag{11}$$

into its phosphate ($\boldsymbol{\mu}^{\mathrm{p}}$) and water ($\boldsymbol{\mu}^{\mathrm{w}}$) parts reconstructed from the respective molecular dipoles $\boldsymbol{\mu}_i$. According to Eq. (10) the IR cross-section is determined by a scalar product, which decomposes by Eq. (11) and

$$\hat{\boldsymbol{\mu}}_g^* \hat{\boldsymbol{\mu}}_g = \underbrace{\hat{\boldsymbol{\mu}}_g^{\mathrm{p}*} \hat{\boldsymbol{\mu}}_g^{\mathrm{p}}}_{\text{phosphate}} + \underbrace{\hat{\boldsymbol{\mu}}_g^{\mathrm{w}*} \hat{\boldsymbol{\mu}}_g^{\mathrm{w}}}_{\text{water}} + \underbrace{\hat{\boldsymbol{\mu}}_g^{\mathrm{p}*} \hat{\boldsymbol{\mu}}_g^{\mathrm{w}} + \hat{\boldsymbol{\mu}}_g^{\mathrm{w}*} \hat{\boldsymbol{\mu}}_g^{\mathrm{p}}}_{\text{cross term}} \tag{12}$$

into contributions of the phosphate, of the water, and of cross terms.

Because our BOMD trajectories covered only a few ten ps, the contributions of the water and those of the cross terms turned out to be far from convergence (see Supplementary Information, SI, Figure S5). A substantially larger number of trajectories would have been required for an acceptable convergence of these contributions,[33] which was excluded, however, due to the significant computational effort associated with our quite large system sizes and the required evaluation of the HF exchange for the B3LYP functional. On the other hand, the contributions of the phosphates to the total absorption cross section turned out to show a much better convergence, which was sufficient for a reliable computation of the band positions and rough estimates of intensities.

Vibrational analysis. The bands in the computed spectra were assigned to vibrational modes using the generalized normal coordinate (GNC) analysis proposed by Mathias *et al.*[31,34–36,62] GNC derives approximate normal coordinates by minimizing mutual cross correlations. The projection

of the GNCs onto internal coordinates classifies the GNCs in terms of specific local modes. As a result, the GNC algorithm tessellates the IR spectrum into contributions of these modes and, thus, into assigned IR bands. Because GNC does not assume equipartition, it yields temperatures of the various modes, which can be utilized to correct the band intensities toward the ensemble limit.[36]

To determine solvatochromic shifts we also computed the vibrational spectra of HPO_4^{2-} and $H_2PO_4^-$ *in vacuo* within the harmonic approximation. The Poisson equations for the isolated molecules were solved in cubes of 20 Å side length using the Martyna-Tuckerman solver;[63] all other parameters were chosen equal to the BOMD calculations. The Hessians were obtained by finite differences employing coordinate displacements of ± 0.0025 Å for all atoms.

Results and Discussion

The IR spectrum of $H_2PO_4^-$ in water. We start our vibrational analysis of the phosphate ions in aqueous solution by considering $H_2PO_4^-$. Figure 1 compares its experimental IR spectrum[10] (gray shaded area) with the prediction Eq. (10) derived from the B3LYP-D BOMD trajectories (dotted line). The additional colored lines represent the results of the GNC vibrational analysis (see Methods). Each of these lines indicates an IR band which belongs to a specific vibrational mode. The attached labels code the mode characters, where ν and δ mark stretching and bending modes, subscripts 's' and 'a' signify symmetric and anti-symmetric couplings, and O^H and O^- denote protonated and unprotonated phosphate oxygens, respectively.

Comparing in Figure 1 the B3LYP-D BOMD prediction (dotted line) with the experimental observation (gray area) we see that the spectral positions of the four computed stretching bands $\nu_a(PO^-)$, $\nu_s(PO^-)$, $\nu_a(PO^H)$, and $\nu_a(PO^H)$ almost perfectly match positions of experimental peaks. For the anti-symmetric bending mode $\delta_a(H)$ of the two phosphate hydrogens, however, B3LYP-D BOMD predicts a quite sharp peak at 1241 cm^{-1}, whereas IR spectroscopy finds[10] a broad shoulder centered at 1213 cm^{-1}. The symmetric bending mode $\delta_s(H)$ is calculated at higher frequencies outside the window of the experimental data. The three IR active bending modes

$\delta(\mathrm{OPO})$ of the phosphate core are located by our simulations at frequencies below $600\,\mathrm{cm}^{-1}$, where they nicely cover the experimental peak at $521\,\mathrm{cm}^{-1}$. The indicated assignments to local modes (colored lines) agree with earlier characterizations derived from DFT/MM INMA calculations[10] and, hence, corroborate the reliability of the GNC approach for the identification of vibrational modes from MD simulations.

Note here, that the drawn bands of local stretching and bending modes (colored lines) represent only a subset of all modes. In the depicted spectral range there are two additional modes contributing to the IR intensity in the form of two broad bands. We have omitted these bands, which belong to torsional modes of the hydrogens, from the drawing to avoid an overcrowding of colored lines. As a result, the dotted spectrum does not represent the sum of the depicted colored bands.

While the relative intensities predicted for the stretching bands ν and for the phosphate core bending bands $\delta(\mathrm{OPO})$ still resemble the spectroscopic findings to a certain degree, larger intensity deviations are found for the anti-symmetric hydrogen bend $\delta_{\mathrm{a}}(\mathrm{H})$. An inspection of the trajectories reveals that the involved OH-groups feature strong hydrogen bonds (H-bonds) with the surrounding water and it is most likely that our limited sampling of H-bonded structures missed many of the actually occurring structures. Correspondingly our simulations predict a quite sharp peak at $1241\,\mathrm{cm}^{-1}$, whereas a more extended sampling should yield a much broader $\delta_{\mathrm{a}}(\mathrm{H})$ band. Note that the intensities of the the stretching bands $\nu_{\mathrm{a}}(\mathrm{PO}^{-})$, $\nu_{\mathrm{s}}(\mathrm{PO}^{-})$, $\nu_{\mathrm{a}}(\mathrm{PO^H})$, and $\nu_{\mathrm{a}}(\mathrm{PO^H})$ are satisfactorily described mainly because of the temperature correction[36] employed by GNC (uncorrected data are presented in Figure S6 of the SI).

The above comparison of the simulated with the experimental IR spectra of $H_2PO_4^-$ in water has demonstrated that the spectroscopic positions of the prominent stretching bands are excellently reproduced by the applied (and computationally very expensive) B3LYP-D BOMD approach. This accurate description is achieved despite the quite limited ensemble of solvation structures sampled by the BOMD trajectories, which covered only a total time span of 50 ps. Moreover it has shown that a likewise reliable computation of IR intensities would require much more extended BOMD simulations. For our following comparative evaluation of other computational methods we there-

fore restrict our attention to the vibrational frequencies of the modes analyzed in Figure 1 although our BOMD simulations and GNC vibrational analyses yield complete and fully decomposed IR spectra, of course. Furthermore we will take the frequencies of the PO^{-} and PO^{H} stretching modes, which result from B3LYP-D calculations, as supposedly "exact" reference data that allow us to judge the quality of other methods.

BP/MM INMA versus B3LYP-D BOMD. In our evaluation of computational methods we first compare in Figure 2 the vibrational spectra of (A) $H_2PO_4^-$ and (B) HPO_4^{2-} in aqueous solution (sol.) resulting from our B3LYP-D BOMD simulations (second columns) with the previous BP/MM INMA description[10] (fourth columns) and with the available experimental evidence[10] (central columns). In addition, the figure presents B3LYP-D (first columns) and BP (fifth columns) normal mode analyses of the phosphate ions isolated *in vacuo* (vac.). To enable a fair comparison of the presumably less accurate BP results with the accurate B3LYP-D predictions, we have scaled the BP frequencies by a factor of 1.069. This scaling factor minimizes the root mean square deviation (RMSD) between the PO stretching frequencies resulting from DFT normal mode analyses of isolated $H_2PO_4^-$ molecules (vac.) using the B3LYP-D and BP functionals, respectively.

Considering first in Figure 2A the vibrational frequencies calculated by normal mode analysis for the isolated $H_2PO_4^-$ ion by B3LYP-D (first column) and scaled BP (last column), respectively, one immediately recognizes that the two descriptions are very similar. The PO stretching frequencies differ, after scaling the BP results with 1.069, by a RMSD of only $11\,cm^{-1}$. Also the frequencies of the remaining modes $\delta_a(H)$ and $\delta(OPO)$ are reasonably close.

However, as soon as the frequency shifts are considered that are predicted by B3LYP-D BOMD and BP/MM INMA, respectively, for the transfer of $H_2PO_4^-$ from the gas phase (vac.) into aqueous solution (sol.) the two approaches show striking differences. Whereas the depicted B3LYP-D BOMD frequencies deviate from their experimental counterparts by a RMSD of only $15\,cm^{-1}$, the BP/MM INMA results show a sixfold larger RMSD of $90\,cm^{-1}$. For $H_2PO_4^-$ the close match of the B3LYP-D BOMD and experimental frequencies had become apparent already in Figure 1. Hence,

the important messages of Figure 2A are (i) the sizes and directions of the various solvatochromic frequency shifts, which are predicted by our reference approach B3LYP-D and are visualized in the transition from the first to the second column, and (ii) the alteration of the description of solvatochromy that is caused by simplifying the model of the surrounding water from B3LYP-D to the MM force field TIP3P. For the latter case the solvatochromic shifts are seen in the transition from the fifth to the fourth column. Note that the frequency shifts of the B3LYP-D BOMD approach partially includes effects of anharmonicity, which is probed by the dynamics but is missing in the harmonic INMA approach. For the P–O stretches anharmonicity leads to a general redshift on the order of $10\,\mathrm{cm}^{-1}$ to $20\,\mathrm{cm}^{-1}$, which we have estimated from calculations of the isolated phosphate species $H_{3-n}PO_4^{n-}$, $n = 0,1,2$ (see Supporting Information, Table S1). Since this anharmonicity redshift is much smaller than the RMSD of the BP/MM INMA results from the experimental values and because it is hardly affected by the solvent,[64] it cannot account for the huge deviations of the INMA description.

Encouraged by fact that the B3LYP-D BOMD results excellently reproduce the observed vibrational frequencies of $H_2PO_4^-$ in aqueous solution we may safely assume that the B3LYP-D description of solvatochromic frequency shifts is of similar quality. Thus, according to Figure 2A the two high-frequency PO^- stretches should exhibit sizable red-shifts and the two lower frequency PO^H stretches blue-shifts upon solvation. The latter prediction of blue-shifts is qualitatively retained in the transition BP(vac.) to BP(sol.). However, the DFT/MM hybrid calculation BP(sol.) assigns to the $\nu_a(PO^-)$ mode only a tiny red-shift and to the $\nu_s(PO^-)$ mode even a blue- instead of a red-shift. Similarly, the large solvatochromic blue-shift predicted by B3LYP-D for the $\delta_a(H)$ bending mode is strongly underestimated by the BP(sol.) model and the solvatochromic B3LYP-D red-shift of the $\delta(OPO)$ band is converted by BP(sol.) into a blue-shift.

The differences between the B3LYP-D BOMD and BP/MM INMA descriptions of solvatochromic frequency shifts, which we just have stated for the $H_2PO_4^-$ ion, are transferable to the case of HPO_4^{2-} as becomes apparent by a glance at Figure 2B. Also in this case B3LYP-D (sol.) provides a very good description of the experimental evidence [exp.(sol.)] featuring an RMSD of only

$32\,\mathrm{cm}^{-1}$. Also here the PO^- stretches should be shifted to the red and the PO^H stretch to the blue upon solvation. Also here BP(sol.) qualitatively captures the latter blue shift, but fails on the PO^- stretches, for which it predicts blue- instead of red-shifts. Similar considerations apply to the descriptions of the $\delta(\mathrm{H})$ and $\delta(\mathrm{OPO})$ solvatochromic shifts.

Summarizing we may state that the intra-molecular forces resulting from a *properly scaled* BP description of the isolated phosphate ions are reasonably accurate as judged from comparisons with the B3LYP reference data and that the MM description of the surrounding water by the TIP3P model yields quantitatively and, in part, also qualitatively wrong descriptions of the effects of solvation on the phosphate vibrational spectra.

Note here that the above analysis of the BP/MM shortcomings in the description of solvatochromic shifts substantially differs from the original one,[10] because the comparison of the BP gas-phase vibrational spectra with the B3LYP reference results has now unexpectedly[9,10] revealed that the BP frequencies must be scaled in the specific case of the phosphate ions by the quite large factor of 1.069. This insight is at variance with the original assumption,[10] which had been based on calculations for neutral compounds,[26,65] that a scaling factor of 1.0 should be appropriate. Comparisons of the unscaled BP/MM results with experimental solution spectra had falsely suggested that the solvation effects induced by the TIP3P water environment are too weak to induce the seemingly required strong blue-shifts, whereas the properly scaled BP results now show that the frequencies of some modes should experience blue- and others red-shifts upon solvation. This finding suggests that the ensembles of the TIP3P solvation structures surrounding the phosphates are wrong.

To check the latter hypothesis we have visually inspected the solvation structures used in the INMA calculations of the original investigation.[10] We found that these structures are much less well-ordered than those sampled by our BOMD simulations with the B3LYP-D and BLYP-D functionals. Correspondingly, some of the solvatochromic shifts predicted by the INMA calculations are even reversed with respect to those of the BOMD simulations, since particularly the magnitude and direction of the $\nu(PO^-)$ shifts delicately depend on the local electric fields and hydrogen bond-

ing structures.[64] This indicates that the simplistic TIP3P model, carrying an erroneous quadrupole moment and lacking polarizablity, does not sample the solvation structures correctly and leads to the noted ill-descriptions of solvatochromic frequency shifts. However, for a detailed check of the latter claim DFT/MM calculations are required, in which the aqueous environment is modeled by a more accurate water model. Within the framework of our current study we solely can demonstrate that the qualitative shortcomings of the earlier description are definitely not due to the use of a simple gradient-corrected functional like BP and to the associated neglect of HF exchange. For a proof we will now compare the solvation effects in the vibrational spectra of the phosphates predicted by the comparably simple gradient corrected density functionals BLYP-D and BP using the former in a BOMD and the latter in a DFT/MM INMA setting.

BP/MM INMA versus BLYP-D BOMD. If one simplifies the DFT description by choosing BLYP-D instead of the B3LYP-D reference functional, one may expect that the BLYP-D frequencies have to be scaled for a favorable comparison with experimental data. Like in the case of BP we chose the PO stretching frequencies calculated by normal mode analysis for the isolated $H_2PO_4^-$ for computing a scaling factor which minimizes the RMSD of the BLYP-D from the corresponding B3LYP-D results. We found the factor of 1.032, which is smaller than the BP factor of 1.069 but is also larger than one. The resulting RMSD is $6\,cm^{-1}$ indicating that B3LYP-D and scaled BLYP-D describe the PO stretching frequencies of the isolated $H_2PO_4^-$ ion at a similar quality. Here, scaled BLYP-D and scaled BP deviate by an RMSD of $11\,cm^{-1}$ proving the expected near equivalence of the two gradient corrected functionals.

Figure 3 compares the band positions of the $H_2PO_4^-$ species obtained by scaled BLYP-D with the experimental data[10] and with the scaled BP results. Comparing first the columns BLYP-D (vac.) and BP(vac.) underlines the near equivalence of these two scaled descriptions. Next, the scaled PO frequencies in column BLYP-D (sol.), which result from our BOMD simulation, are seen to match the experimental data [exp.(sol.)] qualitatively well. As expected, the PO^- stretches are shifted to the red and the PO^H stretches to the blue upon solvation. Quantitatively, however, the

BLYP-D (sol.) description of the solution data exp.(sol.) is substantially worse than that noted for B3LYP-D (sol.) above as is witnessed by the much larger RMSD of $40\,\mathrm{cm}^{-1}$. On the other hand, it is still substantially better than the DFT/MM hybrid model BP/MM (sol.) with its two times larger RMSD and its qualitatively wrong solvatochromic shifts.

Thus, our above hypothesis that the qualitative shortcomings of the DFT/MM hybrid model BP(sol.) in the description of the $H_2PO_4^-$ solvatochromic frequency shifts are mainly due to insufficiencies of the MM force field and cannot be attributed to the lacking HF exchange has been proven. On the other hand, the much larger RMSD between the scaled BLYP-D BOMD frequencies and the experimental data shows that the enhanced computational effort connected with the B3LYP-D inclusion of HF exchange actually has paid off. Despite the limited sampling the B3LYP-D BOMD simulation has rendered an excellent description of the observed IR spectrum particularly for the singly charged ion $H_2PO_4^-$ in aqueous solution.

Beyond vibrational spectra our "first principles" BOMD simulations can shed light also on other properties of $H_2PO_4^-$ and HPO_4^{2-} ions in aqueous solution. In the remainder of this section we will concentrate on the structure of the first solvation shell as predicted by B3LYP-D for $H_2PO_4^-$.

We want to emphasize, however, that the Supplementary Information (SI) documents two additional properties that may be of interest to certain readers. These are (i) the density of electronic states obtained for the two ions, which shows that the protonation state of phosphate ions in solution should be discernible by photo-electron spectroscopy (see Figure S8),[66] and (ii) the amounts of charge transfer from the phosphates to the solvent as quantified by atomic Mulliken charges (see Figure S9). This charge transfer is shown to become smaller upon inclusion of HF exchange, in particular for the doubly negative ion HPO_4^{2-}.

Phosphate solvation structures from B3LYP-D BOMD. Figure 4 shows O$\cdots$O (top) and O$\cdots$H (bottom) radial distribution functions (RDFs) $g(r)$ characterizing for $H_2PO_4^-$ the distribution and hydrogen bonding patterns of the surrounding water molecules. RDFs are depicted for the chemically different hydrogen bonding partners occurring in this solute-solvent system which include

the water oxygens O^w and hydrogens H^w, the unprotonated and protonated oxygens O^- and O^H, respectively, of the phosphate as well as the phosphate protons H. For comparison also the RDFs characterizing the water-water interactions are given (blue curves).

Considering first the $O^- \cdots O^w$ RDF represented by the red curve in the top panel one recognizes that the height (2.2) of its first peak at 2.78 Å is smaller than that (2.9) of the water RDF at 2.80 Å. The reason is that the phosphate oxygens O^- offer only three available sites for neighboring water molecules instead of the four tetrahedrally coordinated sites in bulk water. The slightly smaller $O \cdots O$ distances found for the H-bonds between the water molecules and the phosphate O^- indicate that these bonds are a little stronger than those in bulk water. For HPO_4^{2-} (see Figure S7) the first peak moves to the even shorter distance of 2.75 Å reflecting the correspondingly stronger interaction of the water molecules with this doubly charged ion. The first peak of the $O^H \cdots O^w$ RDF (green curve) at 2.73 Å is smaller than those of the other two $O \cdots O$ RDFs indicating a reduced occupancy of H-bonding sites near the protonated phosphate oxygens O^H.

For a more detailed understanding of the latter finding we first note that the phosphate O^H atoms, in contrast to the O^- atoms, can act both as acceptors and as donors of H-bonds. The occupancies and H-bonding strengths of these two different sites are distinguished in the bottom panel of Figure 4 by the RDFs drawn in green and in purple for the acceptor motif $O^H \cdots H^w$ and the donor motif $O^w \cdots H$, respectively. The first peak of the green curve is very small and located at the comparatively large distance of at 1.95 Å indicating that the O^H atoms of $H_2PO_4^-$ accept only weak and sparsely populated H-bonds from the solvent. In contrast, the O^H atoms donate quite strong and well-populated H-bonds to the solvent as demonstrated by the pronounced first peak of the purple curve at 1.78 Å. The integrals over the first peaks of the green and purple RDFs indicate that each acceptor site of the O^H atoms was occupied by only 37 % whereas each donor site was always fully occupied during our B3LYP BOMD simulations.

For completeness we note that each unprotonated phosphate oxygen O^- of $H_2PO_4^-$ can accept three H-bonds from the solvent, each of which showed an occupancy of 91 % (from the red curve, bottom panel). This occupancy is comparable to the one found for the two acceptor sites in bulk

water (93 %; from the blue curve, bottom panel). Like in the top panel the red peak is found at a smaller distance than the blue peak underlining the stated enhanced strengths of the $O^{-}\cdots O^{w}$ H-bonds.

Summary and Outlook

As the most important result of our study we first note that B3LYP-D BOMD simulations of $H_2PO_4^-$ in aqueous solution were capable to reproduce — without any frequency scaling — the observed[10] spectral locations of prominent mid-IR bands (cf. Figure 2) with an RMSD of only $15\,cm^{-1}$. This excellent description of the spectroscopic findings was achieved although the applied BOMD sampling spanned only 55 ps, which is too short to get a statistically converged ensemble of solvent-solute structures and, therefore, prevented a likewise accurate computation of the IR intensities. While being relatively short compared to some relevant physical processes, this sampling time was nevertheless relatively extended, if one considers the significant computational effort posed by these simulations. For the doubly charged anion HPO_4^{2-} the B3LYP-D BOMD description of the IR spectroscopic evidence became a little less accurate as witnessed by the RMSD of $32\,cm^{-1}$.

Dropping HF exchange and choosing the computationally much less demanding GGA in the form of the BLYP-D functional enforced a frequency scaling by the quite large factor of 1.032 to obtain for isolated phosphates a good match with the B3LYP-D reference frequencies. For the BP approach an even larger scaling factor of 1.069 had to be chosen to achieve a match of similar quality. After scaling the BLYP-D BOMD approach furnished for $H_2PO_4^-$ a description of the spectroscopically determined P–O stretching frequencies, which showed the quite large RMSD of $40\,cm^{-1}$, demonstrating that the inclusion of HF exchange is necessary for quantitative DFT descriptions of phosphate IR spectra. Qualitatively, however, BLYP-D BOMD was capable to reproduce the pattern of solvatochromic frequency shifts, which had been reliably identified by B3LYP-D BOMD. In contrast, the previous BP/MM INMA approach[10] — after proper scaling —

turned out to predict different and, thus, wrong patterns of solvatochromic shifts.

For future DFT/MM calculations on vibrational spectra of phosphates in more complex condensed phase environments like proteins, which cannot be tackled by "first principles" BOMD in a foreseeable future, these findings imply that (i) high quality MM models for the electrostatic properties (multipole moments, polarizability) of the surrounding molecules and (ii) the inclusion of HF exchange are mandatory for qualitatively correct and quantitatively accurate descriptions, respectively. The solvation structures sampled by our B3LYP-D BOMD simulation of $H_2PO_4^-$ in aqueous solution and characterized by the RDFs depicted in Figure 4 can then serve as a reference to judge the quality of DFT/MM models for phosphate solvation.

Finally we would like to emphasize once again that we have added further results of our phosphate-water BOMD simulations, which are outside the central scope of this paper, because they are not directly related to solvatochromic effects in the IR spectra of phosphates, into the SI.

Acknowledgement

PT and GM gratefully acknowledge financial support by the Sonderforschungsbereich 749 "Dynamics and Intermediates of Molecular Transformations" (DFG/SFB749-C4) of the Deutsche Forschungsgemeinschaft. JV acknowledges a generous allocation of computer time from the Swiss National Supercomputing Centre (CSCS).

Supporting Information Available

The SI provides five figures (S5-S9) and associated pieces of text describing complementary data on the splitting of the IR absorptions into contributions of phosphates and water [see Eq. (12)], the IR intensity correction by scaling the absorption peaks with the mode temperatures, and the RDFs of the B3LYP-D simulation of HPO_4^{2-}. In addition, the SI explains and documents two observations not discussed in the main manuscript, which are (i) the B3LYP-D prediction on the density of electron states in the phosphate-water systems, which could be interesting in the context of photoelectron spectroscopy, and (ii) the BOMD charge transfer from the phosphate to the surrounding

water, which may be helpful to evaluate the effects of including HF exchange in DFT functionals. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Vetter, I. R.; Wittinghofer, A. *Science* **2001**, *294*, 1299–1304.

(2) Takeuchi, H.; Murata, H.; Harada, I. *J. Am. Chem. Soc.* **1988**, *110*, 392–397.

(3) Barth, A.; Hauser, K.; Mantele, W.; Corrie, J.; Trentham, D. *J. Am. Chem. Soc.* **1995**, *117*, 10311–10316.

(4) Barth, A. *J. Biol. Chem.* **1999**, *274*, 22170–22175.

(5) Allin, C.; Ahmadian, M.; Wittinghofer, A.; Gerwert, K. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 7754–7759.

(6) Karjalainen, E.; Hardell, A.; Barth, A. *Biopolymers* **2006**, *91*, 2282–2289.

(7) Kötting, C.; Blessenohl, M.; Suveyzdis, Y.; Goody, R. S.; Wittinghofer, A.; Gerwert, K. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13911–13916.

(8) Kötting, C.; Kallenbach, A.; Suveyzclis, Y.; Wittinghofer, A.; Gerwert, K. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6260–6265.

(9) Schmitz, M.; Tavan, P. In *Modern methods for theoretical physical chemistry of biopolymers*; Tanaka, S., Lewis, J., Eds.; Elsevier: Amsterdam, 2006; Chapter 8, pp 157–177.

(10) Klähn, M.; Mathias, G.; Kötting, C.; Schlitter, J.; Nonella, M.; Gerwert, K.; Tavan, P. *J. Phys. Chem. A* **2004**, *108*, 6186–6194.

(11) Klähn, M.; Schlitter, J.; Gerwert, K. *Biophys. J.* **2005**, *88*, 3829–3844.

(12) Heesen, H.; Gerwert, K.; Schlitter, J. *FEBS Lett.* **2007**, *581*, 5677–5684.

(13) Nonella, M.; Mathias, G.; Tavan, P. *J. Phys. Chem. A* **2003**, *107*, 8638–8647.

(14) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12233–12246.

(15) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12247–12258.

(16) Nonella, M.; Mathias, G.; Eichinger, M.; Tavan, P. *J. Phys. Chem. B* **2003**, *107*, 316–322.

(17) Babitzki, G.; Mathias, G.; Tavan, P. *J. Phys. Chem. B* **2009**, *113*, 10496–10508.

(18) Rieff, B.; Mathias, G.; Bauer, S.; Tavan, P. *Photochem. Photobiol.* **2011**, *87*, 511–523.

(19) Rieff, B.; Bauer, S.; Mathias, G.; Tavan, P. *J. Phys. Chem. B* **2011**, *115*, 2117–2123.

(20) Rieff, B.; Bauer, S.; Mathias, G.; Tavan, P. *J. Phys. Chem. B* **2011**, *115*, 11239–11253.

(21) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

(22) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(23) Perdew, J.; Yue, W. *Phys. Rev. B* **1986**, *33*, 8800–8802.

(24) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(25) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(26) Neugebauer, J.; Hess, B. A. *J. Chem. Phys.* **2003**, *118,*, 7215–7225.

(27) Gaigeot, M.-P.; Sprik, M. *J. Phys. Chem. B* **2003**, *107*, 10344–10358.

(28) Gaigeot, M.; Vuilleumier, R.; Sprik, M.; Borgis, D. *J. Chem. Theory Comput.* **2005**, *1*, 772–789.

(29) Mathias, G.; Marx, D. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 6980–6985.

(30) Mallik, B.; Semparithi, A.; Chandra, A. *JPCA* **2008**, *112*, 5104–5112.

(31) Ivanov, S. D.; Asvany, O.; Witt, A.; Hugo, E.; Mathias, G.; Redlich, B.; Marx, D.; Schlemmer, S. *Nat. Chem.* **2010**, *2*, 298–302.

(32) Heyden, M.; Sun, J.; Funkner, S.; Mathias, G.; Forbert, H.; Havenith, M.; Marx, D. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 12068–12073.

(33) Sun, J.; Bousquet, D.; Forbert, H.; Marx, D. *J. Chem. Phys.* **2010**, *133*, 114508.

(34) Baer, M.; Marx, D.; Mathias, G. *Angew. Chem. Int. Ed.* **2010**, *49*, 7346–7349.

(35) Baer, M.; Marx, D.; Mathias, G. *ChemPhysChem* **2011**, *12*, 1906–1915.

(36) Mathias, G.; Baer, M. *J. Chem. Theory Comput.* **2011**, *7*, 2028–2039.

(37) Hesske, H.; Urakawa, A.; VandeVondele, J.; Baiker, A. *J. Phys. Chem. C* **2010**, *114*, 15042–15048.

(38) Muniz-Miranda, F.; Pagliai, M.; Cardini, G.; Schettino, V. *J. Chem. Theory Comput.* **2011**, *7*, 1109–1118.

(39) Morales, C.; Thompson, W. *JPCB* **2011**, *115*, 7597–7605.

(40) Mathias, G.; Egwolf, B.; Nonella, M.; Tavan, P. *J. Chem. Phys.* **2003**, *118*, 10847–10860.

(41) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(42) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103–128.

(43) Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703–1710.

(44) Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **1998**, *58*, 3641–3662.

(45) Lippert, G.; Hutter, J.; Parrinello, M. *Mol. Phys.* **1997**, *92*, 477–487.

(46) Guidon, M.; Schiffmann, F.; Hutter, J.; VandeVondele, J. *J. Chem. Phys.* **2008**, *128*, 214104.

(47) Guidon, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2009**, *5*, 3010–3021.

(48) Guidon, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2010**, *6*, 2348–2364.

(49) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(50) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(51) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(52) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(53) VandeVondele, J.; Mohamed, F.; Krack, M.; Hutter, J.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **2005**, *122*, 014515.

(54) Schmidt, J.; VandeVondele, J.; Kuo, I.-F. W.; Sebastiani, D.; Siepmann, J. I.; Hutter, J.; Mundy, C. J. *J. Phys. Chem. B* **2009**, *113*, 11959–11964.

(55) VandeVondele, J.; Hutter, J. *J. Chem. Phys.* **2003**, *118*, 4365–4369.

(56) Martyna, G. J.; Klein, M. L.; Tuckerman, M. E. *J. Chem. Phys.* **1992**, *97*, 2635–2643.

(57) Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: Cambridge, 2009; pp 309–349.

(58) McQuarrie, D. A. *Statistical Mechanics*, 1st ed.; University Science Books, 2000.

(59) Ramírez, R.; López-Ciudad, T.; Kumar, P. P.; Marx, D. *J. Chem. Phys.* **2004**, *121*, 3973–3981.

(60) Egorov, S. A.; Everitt, K. F.; Skinner, J. L. *J. Phys. Chem. A* **1999**, *103*, 9494–9499.

(61) Lawrence, C. P.; Skinner, J. L. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6720–6725.

(62) Mathias, G.; Ivanov, S. D.; Witt, A.; Baer, M. D.; Marx, D. *J. Chem. Theory Comput.* **2012**, *8*, 224–234.

(63) Martyna, G.; Tuckerman, M. *J. Chem. Phys.* **1999**, *110*, 2810–2821.

(64) Levinson, N.; Bolte, E.; Miller, C.; Corcelli, S.; Boxer, S. *J. Am. Chem. Soc.* **2011**, *133*, 13236–13239.

(65) Nonella, M.; Tavan, P. *Chem. Phys.* **1995**, *199*, 19–32.

(66) Winter, B.; Faubel, M. *Chem. Rev.* **2006**, *106*, 1176–1211.

Figure 1: Total (dotted line) and mode-local (colored lines) IR absorptions of $H_2PO_4^-$ in water derived from B3LYP-D BOMD trajectories are compared with an experimental spectrum[10] (gray shaded area). Computed intensities were scaled to match the experimental maximum at 1077 cm^{-1}.

Figure 2: Vibrational levels obtained by B3LYP-D calculations on (A) $H_2PO_4^-$ and (B) HPO_4^{2-} both isolated (vac.) and in aqueous solution (sol.) are compared with measured[10] [exp.(sol.)] and previously calculated[10] BP/MM INMA and BP harmonic frequencies, which were properly scaled (1.069) for fair comparisons (for details see the text).

Figure 3: Scaled (1.032) frequencies from a BLYP-D normal mode analysis (vac.) and from BLYP-D BOMD simulations (sol.) are compared with experimental evidence [exp.(sol.)] and with properly (1.069) rescaled frequencies[10] from a BP normal mode analysis and from BP INMA hybrid calculations.

Figure 4: O$\cdots$O (top) and O$\cdots$H (bottom) pair correlation functions $g(r)$ calculated from the B3LYP-D trajectories of $H_2PO_4^-$ ion in solution. The atoms O^w and H^w belong to water molecules, O^H and O^- are protonated and unprotonated phosphate oxygens, respectively, and H are phosphate protons.

Figure 1

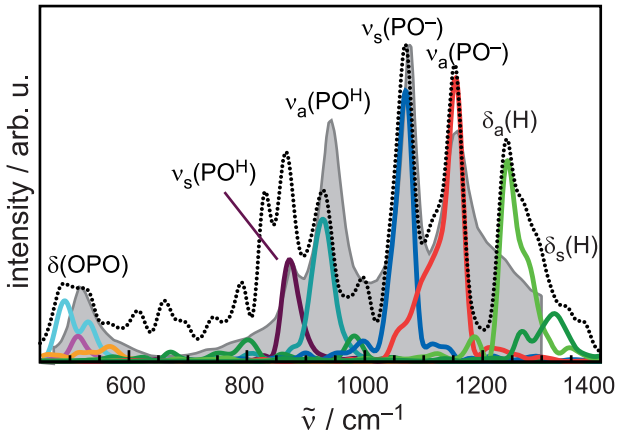

$\nu_s(PO^-)$
$\nu_a(PO^-)$
$\nu_a(PO^H)$
$\delta_a(H)$
$\nu_s(PO^H)$
$\delta_s(H)$
$\delta(OPO)$
intensity / arb. u.
600
800
1000
1200
1400
$\tilde{\nu}$ / cm^{-1}

Figure 2

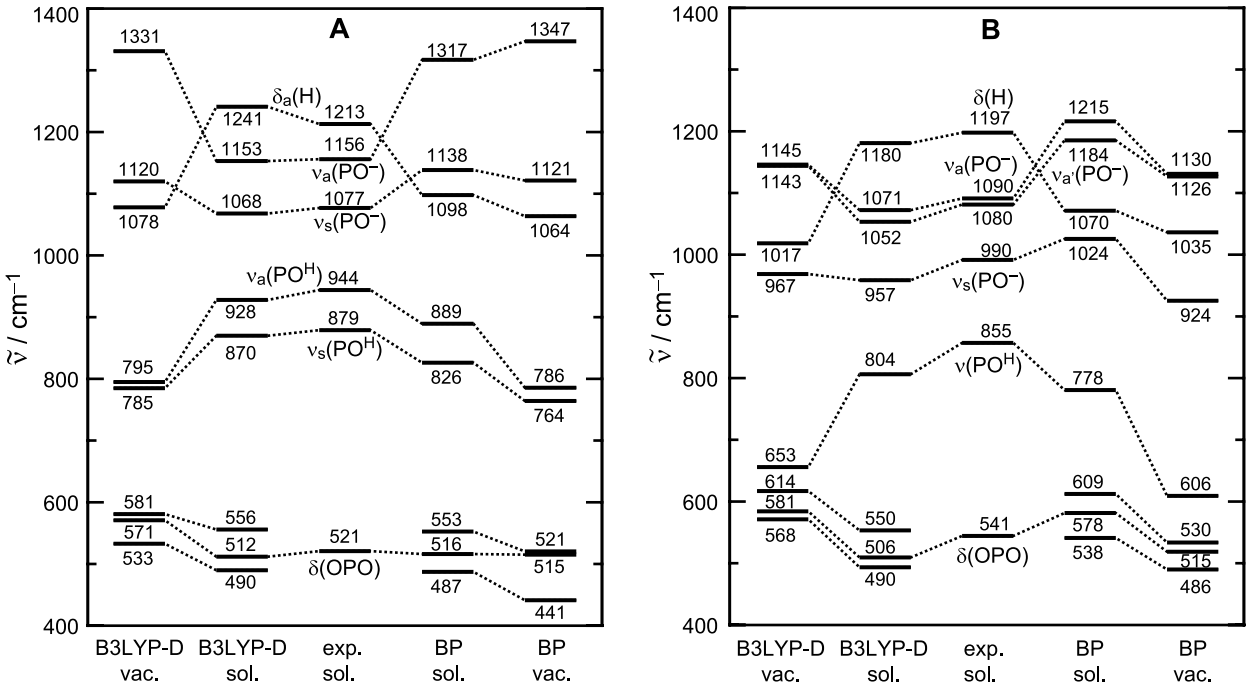

A
B
1400
1200
1000
800
600
400
$\tilde{\nu}$ / cm^{-1}
1331
1241
1153
1120
1078
1068
1213
1156
1077
1317
1347
1138
1121
1098
1064
δ_a(H)
ν_a(PO$^-$)
ν_s(PO$^-$)
ν_a(PO^H)
944
928
879
870
795
785
889
826
786
764
ν_s(PO^H)
581
571
533
556
512
490
521
516
487
553
521
515
441
δ(OPO)
1145
1143
1180
1071
1052
1197
1090
1080
1215
1184
1070
1130
1126
1035
924
δ(H)
ν_a(PO$^-$)
$\nu_{a'}$(PO$^-$)
1017
967
957
990
1024
ν_s(PO$^-$)
855
804
653
614
581
568
778
606
609
578
538
530
515
486
550
506
490
541
ν(PO^H)
δ(OPO)
B3LYP-D vac.
B3LYP-D sol.
exp. sol.
BP sol.
BP vac.

Figure 3

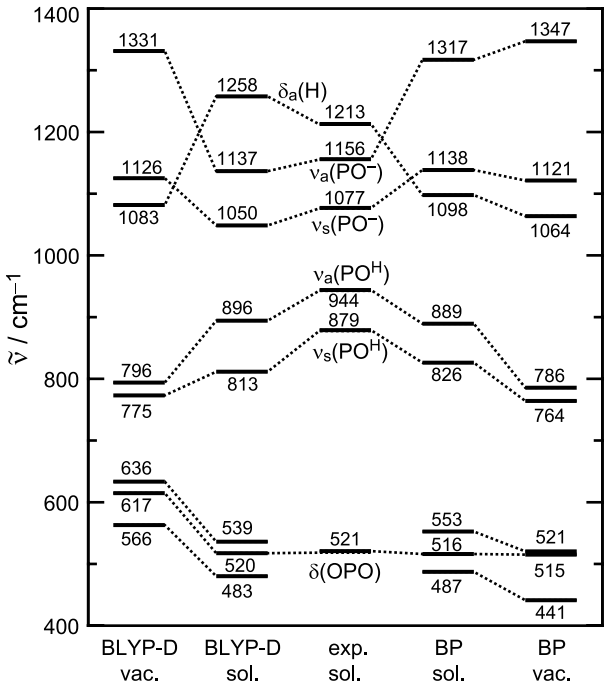
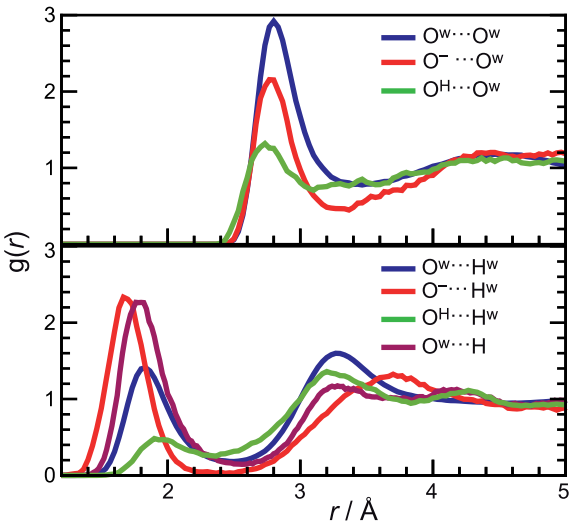
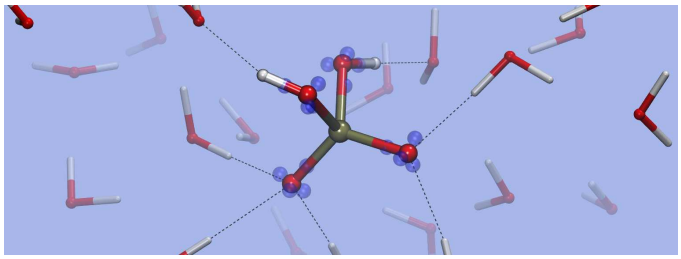

1400
1200
1000
800
600
400
$\tilde{\nu}$ / cm^{-1}
1331
1258
δ_a(H)
1213
1317
1347
1126
1137
1156
ν_a(PO$^-$)
1138
1121
1083
1050
1077
ν_s(PO$^-$)
1098
1064
ν_a(PO^H)
944
896
879
889
ν_s(PO^H)
796
813
826
786
775
764
636
617
566
539
521
553
516
521
520
δ(OPO)
515
483
487
441
BLYP-D vac.
BLYP-D sol.
exp. sol.
BP sol.
BP vac.

Figure 4



Graphical TOC Entry



Phosphate ions play an important role in many biochemical reactions, which can be monitored by vibrational spectroscopy. We demonstrate that first principles MD simulations can compute such spectra quantitatively correct if hybrid density functionals such as B3LYP are employed. The picture shows an $H_2PO_4^-$ ion hydrogen bonded (dashed lines) to the solvent water molecules during the simulation. Blue spheres represent the centers of localized orbitals of the phosphate.